Genome-enabled phylogeographic investigation of the quarantine pathogen *Ralstonia solanacearum* race 3 biovar 2 and screening for sources of resistance against its core effectors

Christopher R. Clarke[1*], David J. Studholme[2*], Byron Hayes[1], Brendan Runde[1], Alexandra Weisberg[1], Rongman Cai[1], Tadeusz Wroblewski[3], Marie Christine Daunay[4], Emmanuel Wicker[5], Jose A. Castillo[6], Boris A. Vinatzer[1]

[1] Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Latham Hall, Ag Quad Lane, Blacksburg, VA, USA;

[2] Biosciences, University of Exeter, Exeter, Devon, UK

[3] Genome Center and Department of Plant Sciences, University of California, Davis, California 95616;

[4] Unité de Genetique et Amelioration des Fruits et Legumes, INRA, Centre d'Avignon, Montfavet, France;

[5] CIRAD, UMR Peuplements Vegetaux et Bioagresseurs en Milieu Tropical (PVBMT), Saint Pierre, La Reunion, France

[6] PROINPA Foundation, Cochabamba, Bolivia

* The first two authors provided equal contributions

Author for correspondence: Boris A. Vinatzer, Tel: +1 540 231 2126, Email: vinatzer@vt.edu

ABSTRACT

Phylogeographic studies inform about routes of pathogen dissemination and are instrumental for improving import/export controls to limit pathogen spread. Genomes of seventeen isolates of the bacterial wilt and potato brown rot pathogen *Ralstonia solanacearum* race 3 biovar 2 (R3bv2), a select agent in the USA, were thus analyzed to get insight into the phylogeography of this pathogen. Most isolates were found to belong to the same genetic lineage that was probably imported from South America into Africa and Europe. At least one more isolate was imported into Europe in a separate event. Moreover, a conserved repertoire of 31 type III-secreted effector genes was identified in all R3bv2 isolates. These genes are excellent candidates to be targeted with plant resistance genes in breeding programs to develop durable disease resistance. Towards this goal, 17 core effectors were tested in eggplant, tomato, pepper, tobacco, and lettuce for induction of a hypersensitive-like response indicative of recognition by cognate resistance genes. Eight of these effectors triggered a response in one or more plant species. These genotypes may thus harbor resistance genes that could be identified, mapped, cloned and expressed in tomato or potato, for which sources of genetic resistance against R3bv2 are extremely limited.

INTRODUCTION

There is no better disease control than disease prevention. Therefore, keeping pathogens out of an unaffected geographic area or eradicating recently introduced pathogens is paramount. To do this efficiently, it is necessary to understand how pathogens are disseminated. Phylogeography, i.e., inferring past migration events by combining phylogenetic reconstruction with geographic information, is instrumental to gain this understanding. Once a pathogen is introduced into a geographic area, disease control becomes necessary. Unfortunately, many crops are affected by pathogens against which no high-yielding, high-quality cultivars are resistant. In addition, complex genetics and heterozygous status of clonally propagated crop species, such as potato, make it extremely difficult to introduce resistance genes by classical breeding from wild resistant relatives (33). Yet, utilization of resistance genes provides a very effective strategy for crop protection, particularly against bacterial plant pathogens for which chemical control and agricultural practices are often only temporary and ineffective solutions (59).

Today, genome sequencing in combination with biotechnology has the potential to greatly improve diagnostics and global tracking of high-risk pathogens and to accelerate the creation of durable disease resistant cultivars. Here we show progress using such an approach for *Ralstonia solanacearum* (*Rs*), a beta-proteobacterium that enters plant roots from the soil through natural openings and wounds, enters and colonizes the xylem where it blocks flow of water and nutrients, and finally causes wilting and eventual death of the host (13). *Rs* pathogenicity is based on a large repertoire of virulence genes encoding type II-secreted extracellular enzymes and type III-secreted effector (T3E) proteins (23), which in some plant genotypes trigger immunity and thus contribute to host range (51). *Rs* overall has a

very large host range, encompassing more than 54 plant families (29). *Rs* thus constitutes a major biotic constraint in tropical and subtropical areas throughout the world, both on subsistence crops (tomato, pepper, eggplant) and cash crops (potato, banana, ginger) (16).

Rs race 3 biovar 2 (R3bv2) is a sub-group of *Rs* that causes bacterial wilt on many solanaceous species. Because R3bv2 causes potato tubers to rot, the disease is also called "brown rot". Based on phylogeny of the *egl* gene (17), most R3bv2 fall within phylotype IIB sequevar 1 (IIB-1) although a few isolates belong to other phylotype II sequevars (9). Based on multilocus sequence analysis, R3bv2 isolates fall in clade 5 (69). Contrary to most *Rs* strains, which are adapted to hot humid climates, R3bv2 strains are able to cause disease at cool temperatures (43). Consequently, this pathogen constitutes a major problem for subsistence potato farmers in many mountainous regions of Africa and South America. In the Mediterranean region, R3bv2 was first identified in the 1920s, in Egypt, Italy and Spain. From there it probably spread with contaminated early ware potatoes to northern Europe where it started to cause frequent outbreaks in the 1990s after being disseminated via contaminated seed potatoes (25, 32). It is now endemic in many waterways and on the Woody Nightshade (*Solanum dulcemara*) in the Netherlands (58), France (63), the UK (49), Spain (6, 36) and Portugal (11). R3bv2 is currently not present in North America although it was accidentally imported at least twice on geranium cuttings from Kenya and Guatemala, but in both cases it was possible to eradicate it before it became endemic (18, 70). Because of its potentially devastating effect on the US potato industry, R3bv2 is on the USDA – APHIS list of select agents (8, 36) and as such under stringent quarantine regulation.

The limited genetic diversity of R3bv2 strains identified by PCR-RFLP, AFLP, 16S rRNA, and 16S-23S suggests a recent evolutionary origin of R3bv2 (52, 62). However, pulsed field gel electrophoresis and rep-PCR (54), microarrays (9), and Multilocus Variable Number Tandem Repeat Analysis (MLVA) (46) provided enough resolution to distinguish R3bv2 isolates into separate populations, some of which were found to be associated with different geographic regions. Single nucleotide polymorphisms (SNPs) have been successful in reconstructing detailed phylogeographies of several human pathogens, for example of the plague pathogen *Yersinia pestis* (45), because they allow determination of ancestral versus derived relationships. SNPs have also been used with some success in resolving the phylogeography of plant pathogenic bacteria, for example of *Pseudomonas syringae* pv. *actinidiae* (42), but they have not been used yet for phylogeographic studies of *Rs*.

Unfortunately, there is not a single potato cultivar resistant to any *Rs* strain (19). While some wild potato relatives are resistant, for example *S. commersonii* (53), resistance is difficult to introduce into cultivated potato because of different ploidy levels (33). Moreover, R3bv2 has been found virulent and very aggressive on most bacterial wilt-resistant tomato genotypes that contain genes for quantitative resistance (5, 67) (38, 68).

The best-characterized gene that provides resistance to *Rs* is the *Arabidopsis thaliana RRS1-R* gene, which recognizes the T3E PopP2 (14, 15), recently renamed RipP2 (50). In crop species, the only dominant genes for resistance to *Rs* have so far been mapped in eggplant (37, 72) and indication for a resistance gene to an *Rs* effector has also been found in the wild eggplant relative *Solanum torvum* (48). Moreover, Lebeau and colleagues (38) found eggplant germplasm to encompass some remarkable genotypes with resistance to multiple *Rs* strains, including R3bv2.

Because tomato, potato, and eggplant all belong to the same botanical family, expression of disease resistance genes from eggplant in potato or tomato will likely provide resistance, due to compatibility of signal transduction components (31, 61). Therefore, identifying the underlying resistance genes in these eggplant genotypes is paramount.

We have shown in the past that using *Agrobacterium*-mediated transient assays to express individual T3Es in plants and screening for a hypersensitive response (HR) - like plant reaction is an efficient way to detect the presence of potential resistance genes (66, 71). These genes could then be genetically mapped, cloned and expressed in crop cultivars. The problem is that pathogens can overcome a single resistance gene relatively easily by losing or mutating the effector that is being recognized (22, 35). However, Bart and colleagues used the example of *Xanthomonas axonopodis* pv. *manihotis* to show that by sequencing a large number of strains it is possible to identify the core set of T3Es of a pathogen (2). The assumption is that effectors present in most strains of a pathogen play an essential role in pathogen virulence and thus cannot be easily lost by the pathogen after a new resistance gene is employed in the host. Resistance genes that recognize such conserved effectors are thus expected to be more durable than resistance genes that recognize dispensable effectors.

Here, we analyzed the genomes of seventeen R3bv2 isolates collected over 58 years on different continents to get insight into R3bv2 phylogeography and to identify the R3bv2 core effector repertoire to target in crop breeding. We previously tested putative R3bv2 effectors in transient assays on lettuce, tomato, pepper, and tobacco genotypes to survey them for potential resistance genes (71). Here we integrate these results with new results obtained with two eggplant genotypes, one of which is

resistant to R3bv2 (37), to prioritize effectors to target in breeding for bacterial wilt/brown rot resistance

MATERIALS AND METHODS

Bacterial Growth and DNA extraction

Bacterial strains were stored in Cryobank™ beads at -80°C. They were grown first on Nutrient Broth overnight at 28°C, then streaked (50 µL) on Kelman's triphenyl tetrazolium chloride (TZC) agar medium (34) supplemented with 1% yeast extract, and sub-cultured two days at 28°C. DNA was extracted from fresh cultures (~ 1.0 to 2.0 x 10^9 cells), using the QIAamp "DNA Blood and Tissue Kit" (QIAGEN), and checked for titer and quality using a Nanodrop device.

Genome Sequencing

Paired-end Illumina sequencing was performed for 76 cycles on an Illumina GAIIx with PE module at the Center for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario, Canada as previously described (42). 454 sequencing was performed on a GS FLX sequencer with Titanium chemistry and sequencing reads were assembled with Newbler v. 2.5 at the Virginia Bioinformatics Institute, Blacksburg, Virginia, USA. Sequencing details are listed in Table 2.

SNP Identification

To facilitate identification of single-nucleotide polymorphisms, we aligned the sequence reads (454 or Illumina – see Table 2) against the reference genome sequence of *R. solanacearum* strain Po82 (73) using BWA-mem version 0.7.5a-r405 (39) with default parameter values and excluding any reads that did not map uniquely to a

166 single site on the reference genome. From the resulting alignments, we generated
167 pileup files using SAMtools version 0.1.19-96b5f2294a (40). We then parsed the
168 pileup-formatted alignments to examine the polymorphism status of each single-
169 nucleotide site in the entire 5,430,263-bp reference genome. For each single-
170 nucleotide site we categorized it as either ambiguous or unambiguous. A site was
171 considered to be unambiguous only if there was at least 5X coverage by genomic
172 sequence reads from each and every bacterial isolate and only if for each and every
173 bacterial isolate, at least 95% of the aligned reads were in agreement. Any sites that
174 did not satisfy these criteria were considered to be ambiguous and excluded from
175 further analysis. Over the remaining unambiguous sites, we could be very confident in
176 the genotype for all the sequenced isolates.
177 The sequence reads for several genomes that had been previously published by
178 other laboratories were not publicly available at the time of this study. Therefore, we
179 generated 50X coverage of simulated 'pseudo-reads' and aligned these to the
180 reference genome by the same method. The genomes for which we had to resort to
181 using pseudo-reads were the genomes of the R3bv2 isolates UW551 and IPO1609 and
182 the following non-R3bv2 genomes: IPO1609, CMR15, FJAT-1458, FJAT-91,
183 FQY_4, GMI1000, MolK2, NCPPB909, P673, Po82, PSI07, R229, R24, Rs-09-161,
184 Rs-10-244, SD54, UW551 and Y45. The pseudo-reads comprised randomly selected
185 50-bp substrings from the genome sequence; substitution errors were introduced
186 randomly into the pseudo-reads at a rate of 0.1% per base-pair. Clearly there is a
187 significantly higher risk of false-positive variants being called based on these pseudo-
188 reads since the use of pseudo-reads derived from a single consensus genome assembly
189 does not benefit from the alignment of multiple independent sequence reads as is the
190 case for the Illumina and 454 reads. The accuracy of these previously published

genome assemblies (and hence the pseudo-reads derived from them) is limited by the depth and reliability of the Sanger sequence reads used to assemble them and by the accuracy of the assembly process. Therefore, we might expect to observe an elevated number of private variants in these bacterial isolates due to the inevitable presence of errors in their genome sequence assemblies.

Construction of phylogenetic trees

We used the results of the genome-wide SNP-calling to investigate the phylogenetic relationships among the sequenced bacterial isolates. The SNP sites were concatenated together to generate a multiple sequence alignment, on which was based a maximum likelihood (ML) phylogenetic tree, built using RAxML (57). Twenty ML searches were performed for each along with 1200 non-parametric bootstrap replicates using the ASC_GTRGAMMA model. Bootstrap values as a percentage of 1200 replicates were mapped to branches on the best ML tree, and branches with less than 50% bootstrap support were collapsed into polytomies using TreeCollapseCL4 (30).

Identification of effector repertoires

We checked for presence of each of the 60 predicted effector genes previously catalogued (50) in *Rs* UW551 (21) by aligning sequence reads against the gene sequences using BWA-mem version 0.7.5a-r405 (39) with default parameter values (allowing non-uniquely mapping reads to match multiple genomic sites). From the resulting alignments, we generated pileup files using SAMtools version 0.1.19-96b5f2294a (40) and parsed these pileup files to determine the proportion of the gene that was covered by aligned sequence reads (that is, the breadth of coverage for each

gene). To check for frame-shifts and premature stop-codons, we generated consensus sequences for each gene for each sequenced isolate (based on the pileup files). We then checked whether each consensus effector gene sequence by translating from nucleotide into amino acid sequence using BioPerl (56).

Agrobacterium-*mediated transient expression of effectors*

Eggplant seeds were sowed in Sunshine 1 Soil mix (company, location) and grown at 22ºC on a 16 hour day cycle in Conviron (Henderson, North Carolina) PGC20 chamber. Other plant species and cultivars tested were grown in the greenhouse condition as described previously (71). Five to seven weeks old eggplants, two to three weeks old lettuce plants, and four to six weeks old tomato, pepper and *N. benthamiana* plants were used for transient assays.

Thirty-two effectors from *R. solanacearum* strain UW551 that were previously cloned into the vector pBAV150 (66) under control of a dexamethasone-inducible promoter were transformed via conjugation into *Agrobacterium tumefaciens* strain C58 (28). *Agrobacterium*-mediated transient expression was performed similar to previously described protocols (71). Briefly, *Agrobacterium* was inoculated into liquid Luria-Bertani Broth with antibiotics and 200uM Acetosyringone and incubated at 28ºC at 200RPM for 18 hours in 50ml conical tubes. Cells were harvested via centrifugation at 4000 RPM. Cells were resuspended in sterile H_2O and centrifuged to pellet a second time. Cells were resuspended in sterile H_2O again and adjusted to an approximate optical density at 600nm (OD_{600}) of 0.5-1 and acetosyringone was added to a final concentration of 200uM and incubated at 28ºC at 200RPM for an additional 4-6 hours. The OD_{600} was adjusted to 0.3 and the adaxial side of leaves were infiltrated with a 1ml blunt-end syringe. Two days after infiltration, infiltrated

leaves were sprayed with 30uM Dexamethasone. Five days after infiltration, the severity of induced HR-like symptoms was scored.

RESULTS

A genome-based phylogeny and estimation of the time since the most recent common ancestor for R3bv2

To investigate phylogeny and phylogeography of R3bv2, we chose to sequence the genomes of fourteen R3bv2 isolates (Table 1). The isolates were chosen to maximize the time span of isolation (from 1950 to 2008), geographic location (South America, Africa, Europe, and Asia), and genetic diversity based on previously obtained microarray analysis (9). The publicly available genomes of the R3bv2 isolates UW551 (21), IPO1609 (27), and NCPPB909 (GenBank: JNGD00000000.1, unpublished) were also included in the analysis. Details of the employed sequencing technology and the obtained genome coverage for each genome are listed in Table 2.

To start investigating the phylogeny of R3bv2 strains, we first placed them into a larger phylogenetic context by building a phylogenetic tree including publicly available representatives of the entire *Rs* species complex. The genome of strain Po82 isolated from potato but also pathogenic to other Solanaceae and to banana (73) was used as reference since Po82 is the strain most closely related to R3bv2 whose genome sequence is complete (rather than of draft quality). A total of 1,108,724 base-pairs (bp) of this genome were covered at least 5 times by sequencing reads with consistent base calls (or assembled genome sequences) from each of the sequenced genomes. Therefore, the tree (Figure 1A) was built based on this 20% of the genome for which the genotype could be unambiguously determined for all sequenced isolates. As expected, all newly sequenced R3bv2 isolates cluster closely together in a

clade shared with previously sequenced known members of phylotype IIB: the R3bv2 isolates UW551 (21), IPO1609 (27), and NCPPB909, strain Molk2 (causative agent of banana Moko disease), strain P673 isolated from the ornamental plant *Epipremnum aureum* (pothos) and also able to infect tomato and, to a lesser degree, potato at 18°C (4) and the reference strain Po82 (73).

Next, we built a tree of only R3bv2 genomes. For this tree, 2,006,873 positions were unambiguous (37% of the total Po82 genome), providing greater phylogenetic resolution than the previous tree. Moreover, SNPs were distributed quite evenly over both the chromosome and the megaplasmid (17,852 SNPs within the main chromosome and 7,136 SNPs SNPs within the mega plasmid). Importantly, the tree built on all SNPs (Figure 1B) and the trees built on SNPs from only the main chromosome had the same branching order and the tree built only on SNPs from the mega plasmid had a very similar topology (see Figure S1A and S1B) showing that the main chromosome and megaplasmid have similar evolutionary histories and, thus, the megaplasmid was not frequently exchanged between strains.

The South American isolates 23-10BR, NCPPB282 and Rs2 (isolated in Brazil, Colombia, and Bolivia, respectively), and the Dutch isolate CFBP3858 are on relatively long branches and are distinct from all other R3bv2 isolates, which instead form a tight cluster of highly similar isolates (See supplementary Table S1 for the exact number of pairwise SNPs for all R3bv2 genomes). Within this cluster of 13 isolates, 11 isolates collected in Europe, Africa, and Asia only differ at one to seven SNP loci from each other although they were collected over a time period of 50 years. We thus used the two most similar isolates collected the furthest apart in time (CFBP4808 isolated in 1954 and Pa1 isolated in 2004) to infer a minimal yearly mutation rate by dividing the number of mutations that distinguished these two

genomes by the difference in years of isolation divided by two (since both strains mutated subsequent to diverging from their most recent common ancestor, MRCA). Because only four mutations distinguished the two isolates over the entire length of the 2,006,873 bp that were compared, the inferred yearly mutation rate was 1.99 x 10^{-8} mutations bp^{-1} $year^{-1}$. Applying this rate to the two most divergent R3bv2 strains, 23-10BR and CFBP3858, which are separated by 5370 SNPs - the MRCA of R3bv2 was inferred to have possibly existed as long ago as 67,125 years ago.

Inferred Phylogeography of R3bv2

Within the tree of R3bv2 isolates (Figure 1B), all South American isolates (23-10BR, Rs2, and NCPPB282) are on three of the four most basal branches. Therefore, these isolates can be considered more ancestral than all other R3bv2 isolates, which is typical of strains at the geographic origin of a pathogen (24). Moreover, the relatively large genetic distance between the South American isolates (see Supplementary Table 1) is suggestive of relatively high diversity of R3bv2 in South America, also expected for the geographic origin of a pathogen (26, 60). Therefore, the analyzed genome sequences here are in agreement with the previously assumed South American origin of R3bv2 (13) and migration of R3bv2 out of South America to Africa, Europe, and Asia (Figure 2).

The next objective of this study was to infer the direction of dissemination between Europe, Africa, and Asia. Since most strains isolated in Europe, Africa, and Asia were highly similar to each other and are part of a single clade with a bootstrap value of 100 % (Figure 1B), a single ancestor of these isolates may have been exported from South America some time before 1954 (the year of isolation of the earliest strain, CFBP4808) and then exchanged between Europe, Asia, and Africa.

Interestingly, a small number of SNPs within this clade are informative and can provide some indication about the possible order of events during these exchanges (see also Figure 2). First, CFBP4808 (isolated in Israel in 1954) is located on the most basal branch of the clade because at one SNP locus it retains the ancestral (Po82-like) allele whereas all other isolates in the same clade carry the derived allele. CFBP4808 may thus represent a very early derivative of the R3bv2 strain exported from South America and from which most other European, African, and Asian isolates are derived from. Second, the two isolates from Reunion Island (LNPV28.23 and JT516) share alleles at one SNP locus with the Egyptian isolate CFBP4578, possibly suggesting a Mediterranean origin of the Reunion outbreak. Third and finally, the Dutch isolate IPO1609 and the Swedish isolate CFBP3884 share alleles at two SNP loci and may thus be both derived from a recent ancestor imported into Northern Europe.

The European isolate IPO1609 and the Kenyan isolate UW551 (imported into the USA on geranium) are the only two isolates within the major European, African, and Asian clade having more than 100 private single-nucleotide variants. However, there is 100% bootstrap support for their placement within the same clade, strongly suggesting that they share the same common ancestor exported out of South America with the other isolates in the clade. Possible explanations for this high number of variants private to these two isolates are discussed below.

The Dutch isolate CFBP3858 differs at more than 4000 SNP loci from all other European isolates and lies clearly outside of the main European, African, and Asian clade (Figure 1B). Its ancestor was thus most likely imported into Europe independently from the ancestor of all other analyzed European isolates.

R. solanacearum *R3bv2 core effector repertoire*

Repertoires of type III-secreted effectors are known to be important determinants of the host range of plant pathogens including *Rs* (50). We therefore sought to define the core effector repertoire of the R3bv2 group. To do this, we aligned the R3bv2 genomes sequenced here (Table 2) and the previously sequenced genomes of IPO1609 and NCPPB909 against 58 *R. solanacearum* effectors catalogued in the R3bv2 genome UW551 by Peeters and colleagues (50). Results are summarized in Table 3 and detailed coverage information for each effector in each genome is available in Table S2. We considered an effector to be present in a genome when sequencing reads from that genome covered more than 80% of the length of the UW551 ortholog. Using this cut-off, 31 effectors were found to be present in all analyzed R3bv2 genomes and not to have any frame shift or early stop codon in the gene regions covered by reads. We thus defined these 31 effectors as R3bv2 core effector repertoire. However, we may have slightly underestimated the core effector repertoire because some effectors may have less than 80% coverage due to sequence divergence or biased sequencing coverage of the genome and not because of actual absence of genes. On the other hand, some of the effector genes with sequencing coverage between 80% and 100% may contain stop codons or frame-shift in the regions of the effectors not covered by reads and therefore not detected. However, only two (out of 915) of the effector sequences with over 80% coverage had any observable frame-shift mutation or early stop codons in any of the 17 analyzed genomes: RipAX1, which is truncated in NCPPB282, and RipS3, which is truncated in Rs2. We therefore have confidence that most of the 31 core effectors are truly present and functional in the corresponding genomes.

The R3bv2 core effectors were then compared against the effector repertoires predicted for non-R3bv2 sequenced *Rs* genomes at the RalstoT3E database (50) revealing that 11 of the 31 R3bv2 core effectors represent core effectors of the entire *Rs* species complex: RipE1, RipF1, RipG6, RipH1, RipW, RipAB, RipAC, RipAI, RipAN, RipAR, and RipAY. Interestingly, none of the effectors was present exclusively in R3bv2.

Screening for resistance genes against R3bv2 effectors

Agrobacterium-mediated transient expression of T3E proteins in the leaves of plants carrying corresponding resistance gene results in the induction of defense responses usually associated with programmed cell death known also as hypersensitive response (HR) (3). Therefore, 27 R3bv2 effector genes, including 17 of the R3bv2 core effectors, were expressed in two eggplant genotypes, 20 tomato genotypes, 5 pepper genotypes, and 22 lettuce genotypes to determine the possible presence of resistance genes in these plant genotypes. An additional five genes that had been previously predicted to encode effectors but were not identified as such by Peeters and colleagues (50) were also included: Ucd163 (a predicted membrane-bend lytic murein transglycosylase B); Ucd190 (a predicted PopC-like putative type III effector similar to RipAC); Ucd191 (HpaP, a putative type III secretion substrate specificity switch protein which was recently found not to be secreted nor translocated (41); Ucd197 (a putative Type III effector); Ucd203 (a conserved protein of unknown function). All effectors had been cloned from an R3bv2 strain from Bolivia previously (71). While the results for ten of the analyzed effectors for tomato, pepper, *N. benthamiana*, and lettuce were already reported (71) and have been accessible in the CHARGE database

(http://charge.ucdavis.edu/charge_db/chargedb_index.php), these results were analyzed here for the first time in the context of the R3bv2 core effectorome.

The overall severity of the cell death elicited by transient expression of the effectors across many replicates was demarcated into the broad categories of "Strong", "Intermediate" and "None" (see Figure 3 for examples). Specifically, fifteen of the screened effectors elicited an HR-like plant response on at least one genotype tested suggesting the presence of underlying effector-detecting *R* genes (Table 4, see detailed results in Table S3). Eight of these fifteen effectors were core R3bv2 effectors: RipB1, RipE1, RipG6, RipH2, RipX, RipV1, RipV2, and RipAB. Therefore, the genotypes on which these effectors elicited cell death potentially harbor *R* genes capable of possibly providing durable resistance to R3bv2.

Of special interest, six effectors, including five core effectors (RipE1, RipH2, RipV1, RipV2, and RipX), elicited substantially stronger cell death in eggplant genotype MM853 than in genotype MM738 (Table 4). Because MM853 is resistant to R3bv2 whereas MM738 is susceptible (38), the eggplant genotype MM853 possibly has R genes that recognize one or more of these core effectors.

Discussion

Because of its impact on world-wide potato production (16) and because of its status as select agent in the USA (8), R3bv2 has been the focus of many studies. For example, the phylogenetic relationship between R3bv2 and other *Rs* strains has been determined (7, 9, 69), the genomes of a small number of R3bv2 strains have been sequenced (21, 27), and these genome sequences allowed the prediction of virulence genes (50). However, the geographic origin and intercontinental routes of dissemination of R3bv2 are not well known (9) and, even more importantly, the

identification of resistance genes against R3bv2 is still in its infancy (5, 37, 53). Here we made progress in these two areas.

An estimated timescale for R3bv2 emergence and diversification

While yearly mutation rates and divergence times for human pathogen species have recently been inferred, for example (12, 44, 47), our knowledge of divergence times for bacterial crop pathogens is preliminary at best (65) making it impossible to determine the time scale of bacterial crop pathogen evolution.

Mutation rates can only be inferred reliably if there is a statistically significant correlation between years of isolation and the genetic distance from the MRCA (47). Therefore, genomes of a number of isolates collected over many years need to be obtained whereby the collected isolates need to be closely related so that it can be assumed that the collected isolates in fact evolved from each other during the time frame of collection (instead of being derived from a distant ancestor with most of the mutations having occurred long before the isolates were collected, which would make it impossible to infer the number of mutations that accumulated per year).

Surprisingly, eleven closely related R3bv2 isolates collected over a time span of 50 years did not accumulate more than 7 mutations (in the 36% of the genome that could be compared with high confidence). Importantly, isolates collected in similar years were as different from each other as isolates collected 50 years apart (see Table 1 and Table S1). Therefore, a mutation rate could not be inferred based on the accumulation of mutations over time. Instead, a minimal mutation rate based on the most similar isolates collected 50 years apart was inferred and found to be as low as 1.99×10^{-8} bp^{-1} $year^{-1}$. This is lower than the mutation rates recently calculated for human pathogens (mostly in the range from $3x10^{-5}$ to $3x10^{-7}$; (44) and references

439 therein). The reason could be that *Rs* bacteria may grow – and thus mutate - very
440 slowly in soil and surface water populations, which often constitute the inoculum
441 sources of outbreaks (1, 64) and most of the mutations that get fixed during fast
442 pathogen replication during plant infection may constitute a sink rather than a source
443 according to the model of Sokurenko and colleagues (55).
444 However, at this point we cannot exclude the possibility that we
445 underestimated the mutation rate because of ascertainment bias arising from
446 considering only the 36% of the genome for which we could be confident of the
447 genotype in all our sequenced genomes. By definition, this 36% that is represented in
448 all the sequenced genomes represents part of the core genome rather than the variable
449 genome. Examining this portion of the genome has the advantage that it is less likely
450 to have undergone horizontal transfer and thus reflects the 'true' phylogeny in
451 addition to the advantage of high confidence in the nucleotide sequence in all isolates.
452 However, if there is an over-representation of vertically inherited SNPs in the non-
453 core (variable) portion of the genome, then conversely this implies SNPs are under-
454 represented in the core-genome, including the 36% on which we based our
455 phylogenetic analysis.
456 The high number of private variants in the UW551 and IPO1609 genomes
457 may be due to sequencing errors in these two draft-quality genomes that were
458 obtained by relatively low coverage Sanger sequencing (21, 27). This is in contrast to
459 the situation for the genomes that were sequenced to greater depth by next-generation
460 sequencing methods and for which accuracy benefited from consensus of multiple
461 overlapping independent sequencing reads. Alternatively, UW551 and IPO1609 may
462 be derived from outbreaks during which rates of bacterial replication were high.
463

Estimated frequency and direction of R3bv2 dissemination

The phylogeographic studies undertaken with the plague pathogen *Yersinia pestis* are an impressive example of how whole genome sequencing of bacteria allows reconstruction of historic routes of pathogen dissemination (45). While the number of analyzed R3bv2 isolates in our study is much lower than the number of *Y. pestis* isolates used in the above study, we can still make some conclusions by combining the microbial forensics evidence derived from our data with previous results about the origin and dissemination of R3bv2.

R3bv2 is generally assumed to have first emerged in the Andes of South America (13). In particular, French (20) reports that he found biovar 2 in virgin soil in the Amazon basin and he suggests that potatoes brought from the Peruvian highlands and cultivated in these soils centuries ago got infected there and that infected potatoes were brought back later to the Peruvian Highlands leading to the emergence of brown rot. From South America, R3bv2 is assumed to have been exported with potatoes to the rest of the world and molecular data in agreement with this hypothesis were obtained previously by pulsed field gel electrophoresis and rep-PCR (54), microarray analysis (9), and multilocus sequence analysis (69). However, these previous studies were either based on markers that do not allow determination of ancestral versus derived relationships (9, 54) or were based on a small number of loci (69). Here instead we analyzed over one third of the entire length of R3bv2 genomes using SNPs, which are informative of ancestral versus derived relationships. This approach identified much higher diversity of genomes of South American origin compared to genomes of European origin together with the phylogenetic tree in Figure 1B, which clearly shows that all but one of the European, African, and Asian isolates are derived from more ancestral South American isolates, is so far the strongest evidence of a

South American origin of R3bv2. Moreover, our estimated time since the MRCA of R3bv2 (over 50,000 years ago) clearly places the origin of R3bv2 before the domestication of potato and suggests that the ancestor of R3bv2 was a pathogen of wild potatoes or other Solanaceae.

In regard to the phylogeography of R3bv2 outside of South America, Cellier and colleagues (9) had found that R3bv2 strains were clustered geographically based on a phylogeny and population structure. One R3bv2 group was found to be associated with African and the Indian Ocean region, one group with the Mediterranean regions, and one group with Northern Europe. The northern European isolates IPO1609 (Netherlands) and CFBP3884 (Sweden) cluster together also based on SNPs. However, for the rest our data do not support geographic differentiation of populations. In particular, the Dutch isolates Pa1 and KRZ-5 do not share any alleles exclusively with IPO1609 or CFBP3884 supporting a Northern European population. We thus conclude frequent exchange of R3bv2 between regions, which is in agreement with the pulsed field gel electrophoresis and rep-PCR results that also revealed almost identical isolates from Northern Europe, the Mediterranean region, and Africa (54).

As for the temporal order of transfers, it is interesting to note that isolate CFBP4808 was collected in the Mediterranean region in 1954 (the earliest isolate outside of South America) and is at the same time also the most ancestral isolate of the main clade outside of South America. This agrees with the fact that R3bv2 caused outbreaks in the Mediterranean region starting in the 1920s and that R3bv2-contaminated potatoes were imported from that region into the rest of Europe, possibly being the source of the outbreaks observed in northern Europe starting in the 1980s (25, 32). Our data also support a Mediterranean origin of the Reunion Island

population. Since the only Chinese and the only Kenyan isolates sequenced here are also nested within the European clade, either a Mediterranean or Northern European origin of the R3bv2 populations in these geographic areas is likely as well (see Figure 2).

The core effector repertoire of R3bv2

Bart and colleagues suggested that presence of an effector across all isolates may indicate that the effector is essential for the pathogen to cause disease and that resistance genes against such pathogen core effectors may provide durable resistance (2). While this hypothesis still needs to be empirically tested, it is plausible. We thus determined here the core effector repertoire R3bv2 consisting of 31 effectors and the core effector repertoire of *Rs* overall consisting of 13 effectors. These core effectors can now be prioritized to identify corresponding resistance genes for breeding potato, tomato, and pepper cultivars with durable disease resistance. We expect the members of the R3bv2 core effector repertoire to be particularly important to cause disease on the common R3bv2 hosts potato and tomato while the *Rs* core effector repertoire should consist of the basic tool set used by *Rs* to cause disease on all its hosts.

Several plant genotypes contain putative resistance genes against core R3bv2 effectors

Unfortunately, most *Rs* effectors for which sources of resistance had been previously identified are missing from R3bv2, for example: RipAA and RipP1 (51)), RipP2 (15), and Rip36 alias RipAX2 (48). Therefore, the R3bv2 core effectors identified here need to be given priority in identifying new sources of resistance for breeding/engineering resistance to R3bv2.

539 Towards this goal, a collection of cloned R3bv2 effectors was previously used
540 in *Agrobacterium*-mediated expression to identify potential sources of resistance
541 among tomato, pepper, and lettuce accessions and *N. benthamiana* (71). We extended
542 the search here to two eggplant genotypes: MM853 found to be resistant to most *Rs*
543 strains, including R3bv2 strains, and the susceptible MM738 genotype (38).
544 Combining all results, an interesting picture arose. First of all, none of the five
545 tested *Rs* core effectors triggered strong cell death in any of the tested genotypes: two
546 *Rs* core effectors did not trigger any HR-like response in any genotype (RipAY and
547 RipW), two triggered an intermediate response in only one plant genotype each
548 (RipAB in the tomato accession M82 and RipG6 in the tomato accession V121), and
549 RipE1 triggered an intermediate response in two genotypes (the resistant eggplant
550 genotype MM853 and the tomato genotype VF36). This almost complete absence of
551 HR-like responses to *Rs* core effectors suggests that these effectors have been
552 maintained in all *Rs* strains because plants have not evolved resistance against them.
553 Six of the 12 tested R3bv2 core effectors did not trigger any response in any
554 host genotype and two triggered only intermediate responses (RipV2 and RipX).
555 However, RipH2, RipAT, RipV1, and RipBI triggered intermediate or even strong
556 HR-like responses in multiple genotypes of several plant species, revealing that
557 promising sources of resistance may be available for these R3bv2 core effectors (see
558 Table S3 for the complete list).
559 Particularly promising is the result that the R3bv2 core effector RipH2
560 triggered a strong HR-like cell death response in the resistant genotype MM853 but
561 not in the susceptible genotype MM738. RipH2 is a member of the HLK effector
562 family and ripH1-2-3 alleles from the phylotype I strain OE1-1 were demonstrated to
563 be functionally redundant and mildly involved in virulence on tomato and eggplant

564 (10). RipH2 is thus an excellent candidate effector that is possibly recognized by a
565 cognate *R* gene in genotype MM853, which could be mapped in a cross with the
566 susceptible genotype MM738, cloned, and expressed in tomato and/or potato to confer
567 resistance to R3bv2 strains. The multiple lettuce and tomato accessions that also
568 reacted with a strong cell death response to the R3bv2 core effectors RipH2, RipV1,
569 and RipBI may be very promising resistance sources as well.
570 Interestingly, the two effectors that triggered strong cell death responses in the
571 largest number of plant accessions are the non-core effectors RipAT and RipAV. One
572 possible explanation is that these effectors are in fact missing from some R3bv2
573 isolates because ancestors of these isolates were exposed to plants carrying cognate
574 resistance genes and loss of these effectors thus increased fitness on such plants.
575 Finally, the putative effector Ucd197, translocation of which has not been determined,
576 did trigger a strong cell death response in both eggplant accessions suggesting that it
577 may in fact be an effector.
578 In summary, we have found several promising resistance sources for R3bv2,
579 the resistance genes of which could be mapped and cloned. Importantly,
580 *Agrobacterium*-mediated transient expression could be used for mapping the
581 underlying resistance genes in segregating progenies. This would be very useful since
582 employing actual R3bv2 isolates for this purpose would be very difficult in the USA
583 due to the regulatory restrictions and the risk of accidental spread of R3bv2.
584 In conclusion, we have shown here how genome sequencing of multiple
585 isolates of a select agent and transient expression assays of its effector genes can give
586 new insight into pathogen evolution and dissemination and can constitute a tool for
587 identifying new promising sources of disease resistance. Importantly, this research

588 was done without having to actually do any research with the select agent R3bv2 itself
589 avoiding regulatory hurdles and minimizing the risk of accidental pathogen spread.
590
591

Tables

Table 1: R3bv2 strains used in this study with references to papers in which strains were described previously.

Strain Name	sequevar	Year of isolation	Geographic location of isolation	reference
IPO1609[a]	1	1995	Netherlands	(27)
UW551[a]	1	2003	Kenya/Wisconsin	(21)
NCPPB909[a]	1	1961	Egypt	(54)
23-10BR[b]	27	1981	Brazil	this work
NCPPB282[b]	2	1950	Colombia	(9)
POPS2[b]	1	1980	China	this work
Rs2	1	2008	Bolivia	this work
CFBP4808	1	1954	Israel	(9)
CFBP4578	1	1966	Egypt	(9)
CFBP3927	1	1968	Greece	(9)
CFBP3884	1	1984	Sweden	(9)
CFBP3873	1	1992	Belgium	(9)
JT516	1	1993	Reunion	(9)
CFBP3858	25	1995	Netherlands	(9)
LNPV28.23	1	2004	Reunion	(9)
KZR-5[c]	1	2004	Netherlands	(58)
PA1[c]	1	2004	Netherlands	(58)

[a] genome sequences were not obtained as part of this study but were publicly available
[b] DNA of isolate kindly provided by Dr. Allen (University of Wisconsin, USA)
[c] DNA kindly provided by Dr. Stevens (University of Groningen, The Netherlands)

599 Table 2: Genome Sequencing data

Strain	SRA accession numbers	Average Read length (bp)	Number of reads	Genome coverage depth
23-10BR[d]	SRX702422	433.59 [a]	271,396	20 x
NCPPB282[d]	SRX703654	431.73 [a]	290,941	22 x
POPS2[d]	SRX703652	422.94 [a]	230,546	16 x
Rs2	SRX703653	427.76 [a]	277,564	23 x
CFBP4808	SRX701225	2 x 70 [b]	5,065,895[c]	102 x
CFBP4578	SRX701307	2 x 70 [b]	755,467[c]	13 x
CFBP3927	SRX701324	2 x 70 [b]	1,511,314[c]	24 x
CFBP3884	SRX701786	2 x 70 [b]	2,711,618[c]	45 x
CFBP3873	SRX701787	2 x 70 [b]	4,035,125[c]	84 x
JT516	SRX701788	2 x 70 [b]	3,214,532[c]	54 x
CFBP3858	SRX701790	2 x 70 [b]	9,365,297[c]	140 x
LNPV28.23	SRX701791	2 x 70 [b]	6,619,254[c]	108 x
KZR-5	SRX701792	2 x 70 [b]	4,111,225[c]	81 x
PA1	SRX701793	2 x 70 [b]	4,145,998[c]	85 x

600 [a] Genomes of strains 23-10BR, NCPPB282, POPS2 and Rs2 were sequenced by 454.
601 [b] Genomes of these strains were sequenced by Illumina HiSeq, generating paired
602 reads.
603 [c] These numbers refer to the numbers of read-pairs (rather than individual reads).
604 [d] Genome assemblies have been deposited in GenBank under accession numbers
605 JQOI00000000.1, JQSH00000000.1 and JQSI00000000.1.
606

Table 3: Summary of R3bv2 core effector analysis

R3bv2 Core effectors	R3bv2 non-core effectors
RipAB[a], RipAC, RipAD, RipAE, RipAI, RipAN, RipAP, RipAR, RipAY, RipBI, RipB, RipC1, RipE1, RipF1, RipF2, RipG2, RipG3, RipG4, RipG5, RipG6, RipG7, RipH1, RipH2, RipI, RipN, RipS1, RipU, RipV1, RipV2, RipW, RipX	RipA2, RipAA, RipAJ, RipAL, RipAM, RipAO, RipAQ, RipAS, RipAT, RipAV, RipAX1, RipBH, RipD, RipE2, RipJ, RipM, RipO1, RipQ, RipR, RipS7, RipTPS, RipY, RipA4, RipH3, RipS3, RipZ, Ucd_ID197

[a] Core effectors of *Rs* overall are underlined.

Table 4: Elicitation of cell death by transient expression of *Rs* effectors on eggplant (MM738 and MM853), *Nicotiana benthamiana*, tomato (20 genotypes), pepper (5 genotypes), and lettuce (22 genotypes).

Effector – Ral[a]	Effector – Rip[b]	Lettuce[c]	*Nicotiana*[c]	Pepper[c]	Tomato[c]	MM853	MM738
Ral009t	not a Rip	None	None	None	None	None	None
Ral010t	RipD	Inter.	None	None	None	None	None
Ral012t	RipAA	Inter.	Strong	Strong	Inter.	Inter.	Inter.
Ral013t	**RipW**	None	None	None	None	None	None
Ral014t	**RipV2**	None	None	None	Inter.	Inter.	None
Ral015t	**RipAY**	None	None	None	None	None	None
Ral016t	RipO1	None	None	None	None	None	None
Ral017t	RipAQ	None	None	None	None	None	None
Ral018t	**RipG5**	None	None	None	None	None	None
Ral019t	**RipAB**	None	None	None	Inter.	None	None
Ral021t	**RipG7**	None	None	None	None	None	None
Ral022t	**RipH2**	Strong	None	None	Inter.	Strong	None
Ral023t	RipAJ	None	None	None	None	Inter.	None
Ral024t	**RipG4**	None	None	None	None	None	None
Ral025t	RipE2	None	None	None	None	None	None
Ral027t	**RipF2**	None	None	None	None	None	None
Ral028t	RipAT	Strong	None	Strong	Strong	Strong	Strong
Ral030t	**RipG6**	None	None	None	Inter.	None	None
Ral031t	**RipAP**	None	None	None	None	None	None
Ral033t	RipAV	Strong	None	Strong	Strong	None	None
Ral034t	**RipV1**	Strong	None	Inter.	None	Inter.	None
Ral036t	not a RIP	None	None	None	None	None	None
Ral037t	not a Rip	None	None	None	None	None	None
Ral038t	RipAO	None	None	None	None	None	None
Ral040t	**RipBI**	Inter.	None	Inter.	Strong	Inter.	Strong
Ral041t	**RipE1**	None	None	None	Inter.	Inter.	None
Ral042t	RipAS	None	None	None	None	None	None
Ral043t	not a Rip	None	None	None	None	Strong	Strong
Ral044t	**RipG3**	None	None	None	None	None	None
Ral045t	**RipN**	None	None	None	None	None	None
Ral048t	**RipX**	Inter.	None	None	None	Inter.	None
Ral049t	not a Rip	None	Inter.	None	Inter.	None	None

[a] Ral designations are from UC Davis Charge database (http://charge.ucdavis.edu/charge_db/chargedb_index.php).

[b] Bolded effectors are part of the R3bv2 core effectorome.

[c] Intermediate and strong results for tomato, pepper, and lettuce indicate that cell death was observed in at least one of the tested genotypes.

Supplementary Table Legends

Table S1. Number of SNPs distinguishing R3bv2 genomes from each other and from reference strain 23-10BR over 2,259,494bp.

Table S2. Coverage of 58 UW551 effectors in the 17 other analyzed genomes.

Table S3. Detailed results of cell death resulting from *Agrobacterium* transient expression of 33 cloned effector genes in all screened plant genotypes.

Figure Legends

Figure 1: **A.** Maximum likelihood tree of all newly sequenced and a selection of published *R. solanacearum* genomes. The tree is based on SNPs identified by aligning reads against the genome of *R. solanacearum* Po82. The tree is mid-point rooted. Bootstrap values based on 1200 non-parametric replicates are indicated at nodes. Branch lengths are proportional. The clade corresponding to phylotype IIB is highlighted using a higher weight. **B.** Maximum likelihood tree of R3bv2 isolates based on SNPs when aligning reads against the genome of *R. solanacearum* Po82. Bootstrap values based on 1200 non-parametric replicates are indicated at nodes. The tree is rooted on the genome of Po82. Because branch lengths are very different from each other, a cladogram is shown and the number of SNPs is indicated underneath each branch. The clade containing most European/African/Asian isolates is highlighted using a higher weight.

Figure 2: Inferred routes of dissemination of R3bv2 strains. Strain details are listed in Table 1. European isolates are not placed exactly over their country of isolation because of SPACE limitations. Map from Luke Roberts (http://lukeroberts.deviantart.com/) released under a Creative Commons Attribution-Noncommercial 3.0 license.

Figure 3: Examples of the broad categories listed in Table 3 that were used for classification of cell death elicited by agrobacterium-mediated transient expression of effectors in example plant genotypes tested.

Figure S1: Maximum Likelihood tree of R3bv2 isolates based on SNPs identified in the main chromosome only (A) or in the mega plasmid only (B) using the genome of Po82 as reference for SNP calling and as root for tree construction.

References

1. Álvarez, B., López, M. M., and Biosca, E. G. 2008. Survival strategies and pathogenicity of Ralstonia solanacearum phylotype II subjected to prolonged starvation in environmental water microcosms. Microbiology 154(11):3590-3598 doi:10.1099/mic.0.2008/019448-0.

2. Bart, R., Cohn, M., Kassen, A., McCallum, E. J., Shybut, M., Petriello, A., Krasileva, K., Dahlbeck, D., Medina, C., Alicai, T., Kumar, L., Moreira, L. M., Neto, J. R., Verdier, V., Santana, M. A., Kositcharoenkul, N., Vanderschuren, H., Gruissem, W., Bernal, A., and Staskawicz, B. J. 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proceedings of the National Academy of Sciences 109(28):E1972–E1979 doi:10.1073/pnas.1208003109.

3. Bent, A. F., and Mackey, D. 2007. Elicitors, Effectors, and R Genes: The New Paradigm and a Lifetime Supply of Questions. Annual Review of Phytopathology 45(1):399-436 doi:doi:10.1146/annurev.phyto.45.062806.094427.

4. Bocsanczy, A. M., Huguet-Tapia, J. C., and Norman, D. J. 2014. Whole-Genome Sequence of Ralstonia solanacearum P673, a Strain Capable of Infecting Tomato Plants at Low Temperatures. Genome Announcements 2(1) doi:10.1128/genomeA.00106-14.

5. Carmeille, A., Caranta, C., Dintinger, J., Prior, P., Luisetti, J., and Besse, P. 2006. Identification of QTLs for <i>Ralstonia solanacearum</i> race 3-phylotype II resistance in tomato. TAG Theoretical and Applied Genetics 113(1):110-121 doi:10.1007/s00122-006-0277-3.

691 6. Caruso, P., Palomo, J. L., Bertolini, E., Álvarez, B., López, M. M., and Biosca,
692 E. G. 2005. Seasonal Variation of Ralstonia solanacearum Biovar 2
693 Populations in a Spanish River: Recovery of Stressed Cells at Low
694 Temperatures. Applied and Environmental Microbiology 71(1):140-148
695 doi:10.1128/aem.71.1.140-148.2005.
696 7. Castillo, J. A., and Greenberg, J. T. 2006. Evolutionary Dynamics Of
697 Ralstonia Solanacearum. Appl Environ Microbiol 73(4):1225-38.
698 8. CDC, and APHIS. National Select Agent Registry.
699 9. Cellier, G., Remenant, B., Chiroleu, F., Lefeuvre, P., and Prior, P. 2012.
700 Phylogeny and Population Structure of Brown Rot- and Moko Disease-
701 Causing Strains of Ralstonia solanacearum Phylotype II. Applied and
702 Environmental Microbiology 78(7):2367-2375 doi:10.1128/aem.06123-11.
703 10. Chen, L., Shirota, M., Zhang, Y., Kiba, A., Hikichi, Y., and Ohnishi, K. 2014.
704 Involvement of HLK effectors in Ralstonia solanacearum disease development
705 in tomato. Journal of General Plant Pathology 80(1):79-84
706 doi:10.1007/s10327-013-0490-2.
707 11. Cruz, L., Eloy, M., Quirino, F., Oliveira, H., and Tenreiro, R. 2012. Molecular
708 epidemiology of Ralstonia solanacearum strains from plants and
709 environmental sources in Portugal. European Journal of Plant Pathology
710 133(3):687-706 doi:10.1007/s10658-012-9947-y.
711 12. Cui, Y., Yu, C., Yan, Y., Li, D., Li, Y., Jombart, T., Weinert, L. A., Wang, Z.,
712 Guo, Z., Xu, L., Zhang, Y., Zheng, H., Qin, N., Xiao, X., Wu, M., Wang, X.,
713 Zhou, D., Qi, Z., Du, Z., Wu, H., Yang, X., Cao, H., Wang, H., Wang, J., Yao,
714 S., Rakin, A., Li, Y., Falush, D., Balloux, F., Achtman, M., Song, Y., Wang,
715 J., and Yang, R. 2013. Historical variations in mutation rate in an epidemic

716 pathogen, Yersinia pestis. Proceedings of the National Academy of Sciences
717 110(2):577-582 doi:10.1073/pnas.1205750110.
718 13. Denny, T. P. 2006. Plant pathogenic *Ralstonia* species. Pages 573-644. in:
719 Plant-Associated Bacteria Gnanamanickam, ed. Springer, Dordrecht, The
720 Netherlands.
721 14. Deslandes, L., Olivier, J., Theulieres, F., Hirsch, J., Feng, D. X., Bittner-Eddy,
722 P., Beynon, J., and Marco, Y. 2002. Resistance to Ralstonia solanacearum in
723 Arabidopsis thaliana is conferred by the recessive RRS1-R gene, a member of
724 a novel family of resistance genes. Proc Natl Acad Sci U S A 99(4):2404-9
725 doi:10.1073/pnas.032485099
726 032485099 [pii].
727 15. Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M.,
728 Boucher, C., Somssich, I., Genin, S., and Marco, Y. 2003. Physical interaction
729 between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2,
730 a type III effector targeted to the plant nucleus. Proc Natl Acad Sci U S A
731 100(13):8024-9 doi:10.1073/pnas.1230660100
732 1230660100 [pii].
733 16. Elphinstone, J. G. 2005. The current bacterial wilt situation: a global
734 overview. in: Bacterial wilt disease and the *Ralstonia solanacearum* species
735 complex C. Allen, P. Prior, and A. C. Hayward, eds. APS Press, St Paul, MN.
736 17. Fegan, M., and Prior, P. 2005. How complex is the "Ralstonia solanacearum
737 speceis complex". Pages 449-462. in: Bacterial wilt, the disease and the
738 Ralstonia solanacearum species complex C. Allen, P. Prior, and C. Hayward,
739 eds. APS Press, St. Paul, MN.

740 18. Floyd, J., and Michalak, P. S. 2008. New Pest Response Guidelines: Ralstonia
741 solanacearum race 3 biovar 2. U. A. P. E. a. D. Programs, ed.
742 19. Fock, I., Luisetti, J., Collonnier, C., Vedel, F., Ducreux, G., Kodja, H., and
743 Sihachakr, D. 2005. Solanum phureja amd S. stenotomum are sources of
744 resistance to Ralstonia solanacearum for somatic hybrids of potato. in:
745 Bacterial Wilt Disese and The Ralstonia Solanacearum Species Complex C.
746 Allen, P. Prior, and A. C. Hayward, eds. Amer Phytopathol Soc, Washington,
747 D.C.
748 20. French, E. R. 1979. Classification, Distribution and Origin of *Pseudomonas*
749 *solanacearum*. Pages 28-35. in: Planning Conference on the Developments in
750 Control of Potato Bacterila Diseases. International Potato Center, Lima, Peru.
751 21. Gabriel, D. W., Allen, C., Schell, M., Denny, T. P., Greenberg, J. T., Duan, Y.
752 P., Flores-Cruz, Z., Huang, Q., Clifford, J. M., Presting, G., Gonzalez, E. T.,
753 Reddy, J., Elphinstone, J., Swanson, J., Yao, J., Mulholland, V., Liu, L.,
754 Farmerie, W., Patnaikuni, M., Balogh, B., Norman, D., Alvarez, A., Castillo,
755 J. A., Jones, J., Saddler, G., Walunas, T., Zhukov, A., and Mikhailova, N.
756 2006. Identification of open reading frames unique to a select agent: Ralstonia
757 solanacearum race 3 biovar 2. Mol Plant Microbe Interact 19(1):69-79.
758 22. Gassmann, W., Dahlbeck, D., Chesnokova, O., Minsavage, G. V., Jones, J. B.,
759 and Staskawicz, B. J. 2000. Molecular evolution of virulence in natural field
760 strains of Xanthomonas campestris pv. vesicatoria. J Bacteriol 182(24):7053-
761 9.
762 23. Genin, S., and Denny, T. P. 2012. Pathogenomics of the Ralstonia
763 solanacearum Species Complex. Annual Review of Phytopathology 50(1):67-
764 89 doi:doi:10.1146/annurev-phyto-081211-173000.

765 24. Goss, E. M., Tabima, J. F., Cooke, D. E. L., Restrepo, S., Fry, W. E., Forbes,
766 G. A., Fieland, V. J., Cardenas, M., and Grünwald, N. J. 2014. The Irish potato
767 famine pathogen Phytophthora infestans originated in central Mexico rather
768 than the Andes. Proceedings of the National Academy of Sciences
769 111(24):8791-8796 doi:10.1073/pnas.1401884111.
770 25. Grousset, F., Roy, A. S., and Smith, I. M. 1998. Situation of Ralstonia
771 solanacearum in the EPPO region in 1997. EPPO Bulletin 28(1-2):53-63
772 doi:10.1111/j.1365-2338.1998.tb00702.x.
773 26. Grünwald, N. J., and Flier, W. G. 2005. The Biology of Phytophthora
774 infestans at Its Center of Origin*. Annual Review of Phytopathology
775 43(1):171-190 doi:doi:10.1146/annurev.phyto.43.040204.135906.
776 27. Guidot, A., Elbaz, M., Carrere, S., Siri, M. I., Pianzzola, M. J., Prior, P., and
777 Boucher, C. 2009. Specific Genes from the Potato Brown Rot Strains of
778 Ralstonia solanacearum and Their Potential Use for Strain Detection.
779 Phytopathology 99(9):1105-12 doi:10.1094/PHYTO-99-9-1105.
780 28. Hamilton, R. H., and Fall, M. Z. 1971. The loss of tumor-initiating ability in
781 *Agrobacterium tumefaciens* by incubation at high temperature
782 . Experientia 27:229–323.
783 29. Hayward, A. C. 1991. Biology and epidemiology of bacterial wilt caused by
784 Pseudomonas solanacearum. Annu Rev Phytopathol 29:65-87.
785 30. Hodcroft, E. TreeCollapserCL 4.
786 31. Hulbert, S. H., Webb, C. A., Smith, S. M., and Sun, Q. 2001. RESISTANCE
787 GENE COMPLEXES: Evolution and Utilization. Annual Review of
788 Phytopathology 39(1):285-312 doi:doi:10.1146/annurev.phyto.39.1.285.

789 32. Janse, J. D. 1996. Potato brown rot in western Europe - history, present
790 occurence and some remarks on possible origin. EPPO Bulletin(26):17.
791 33. Jansky, S. 2006. Overcoming hybridization barriers in potato. Plant Breeding
792 125(1):1-12 doi:10.1111/j.1439-0523.2006.01178.x.
793 34. Kelman, A. 1954. The relationship of pathogenicity in Pseudomonas
794 solanacearum to colony appearance on tetrazolium medium. Phytopathology
795 44:693-695.
796 35. Kunkeaw, S., Tan, S., and Coaker, G. 2010. Molecular and evolutionary
797 analyses of Pseudomonas syringae pv. tomato race 1. Mol Plant Microbe
798 Interact 23(4):415-24 doi:10.1094/MPMI-23-4-0415.
799 36. Lambert, C. D. 2002. Agricultural Bioterrorism Protection Act of 2002:
800 Possession, use, and transfer of biological agents and toxins; interim final rule.
801 (7 CFR Part 331). Fed. Regist. 67(240):76908-76938.
802 37. Lebeau, A., Gouy, M., Daunay, M. C., Wicker, E., Chiroleu, F., Prior, P.,
803 Frary, A., and Dintinger, J. 2013. Genetic mapping of a major dominant gene
804 for resistance to Ralstonia solanacearum in eggplant. Theoretical and Applied
805 Genetics 126(1):143-158 doi:10.1007/s00122-012-1969-5.
806 38. Lebeau, A., Daunay, M. C., Frary, A., Palloix, A., Wang, J. F., Dintinger, J.,
807 Chiroleu, F., Wicker, E., and Prior, P. 2011. Bacterial Wilt Resistance in
808 Tomato, Pepper, and Eggplant: Genetic Resources Respond to Diverse Strains
809 in the Ralstonia solanacearum Species Complex. Phytopathology 101(1):154-
810 165 doi:10.1094/phyto-02-10-0048.
811 39. Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with
812 Burrows‚ÄìWheeler transform. Bioinformatics 25(14):1754-1760
813 doi:10.1093/bioinformatics/btp324.

814 40. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth,
815 G., Abecasis, G., Durbin, R., and Subgroup, G. P. D. P. 2009. The Sequence
816 Alignment/Map format and SAMtools. Bioinformatics 25(16):2078-2079
817 doi:10.1093/bioinformatics/btp352.
818 41. Lohou, D., Turner, M., Lonjon, F., Cazalé, A.-C., Peeters, N., Genin, S., and
819 Vailleau, F. 2014. HpaP modulates type III effector secretion in Ralstonia
820 solanacearum and harbours a substrate specificity switch domain essential for
821 virulence. Molecular Plant Pathology 15(6):601-614 doi:10.1111/mpp.12119.
822 42. Mazzaglia, A., Studholme, D. J., Taratufolo, M. C., Cai, R., Almeida, N. F.,
823 Goodman, T., Guttman, D. S., Vinatzer, B. A., and Balestra, G. M. 2012.
824 *Pseudomonas syringae* pv. *actinidiae* (PSA) Isolates from Recent Bacterial
825 Canker of Kiwifruit Outbreaks Belong to the Same Genetic Lineage. PLoS
826 ONE 7(5):e36518 doi:10.1371/journal.pone.0036518.
827 43. Milling, A., Meng, F., Denny, T. P., and Allen, C. 2009. Interactions with
828 Hosts at Cool Temperatures, Not Cold Tolerance, Explain the Unique
829 Epidemiology of Ralstonia solanacearum Race 3 Biovar 2. Phytopathology
830 99(10):1127-1134 doi:10.1094/PHYTO-99-10-1127.
831 44. Morelli, G., Didelot, X., Kusecek, B., Schwarz, S., Bahlawane, C., Falush, D.,
832 Suerbaum, S., and Achtman, M. 2010. Microevolution of *Helicobacter pylori*
833 during Prolonged Infection of Single Hosts and within Families. PLoS Genet
834 6(7):e1001036.
835 45. Morelli, G., Song, Y., Mazzoni, C. J., Eppinger, M., Roumagnac, P., Wagner,
836 D. M., Feldkamp, M., Kusecek, B., Vogler, A. J., Li, Y., Cui, Y., Thomson, N.
837 R., Jombart, T., Leblois, R., Lichtner, P., Rahalison, L., Petersen, J. M.,
838 Balloux, F., Keim, P., Wirth, T., Ravel, J., Yang, R., Carniel, E., and

839 Achtman, M. 2010. *Yersinia pestis* genome sequencing identifies patterns of
840 global phylogenetic diversity. Nat Genet 42(12):1140-3
841 doi:http://www.nature.com/ng/journal/vaop/ncurrent/abs/ng.705.html -
842 supplementary-information.
843 46. N'Guessan, C. A., Brisse, S., Le Roux-Nio, A.-C., Poussier, S., Koné, D., and
844 Wicker, E. 2013. Development of variable number of tandem repeats typing
845 schemes for *Ralstonia solanacearum*, the agent of bacterial wilt, banana Moko
846 disease and potato brown rot. Journal of Microbiological Methods 92(3):366-
847 374 doi:http://dx.doi.org/10.1016/j.mimet.2013.01.012.
848 47. N√ºbel, U., Dordel, J., Kurt, K., Strommenger, B., Westh, H., Shukla, S. K.,
849 ≈Ωemli*ƒ*çkov√°, H., Leblois, R. l., Wirth, T., Jombart, T., Balloux, F. ß., and
850 Witte, W. 2010. A Timescale for Evolution, Population Expansion, and Spatial
851 Spread of an Emerging Clone of Methicillin-Resistant <italic>Staphylococcus
852 aureus</italic>. PLoS Pathog 6(4):e1000855.
853 48. Nahar, K., Matsumoto, I., Taguchi, F., Inagaki, Y., Yamamoto, M., Toyoda,
854 K., Shiraishi, T., Ichinose, Y., and Mukaihara, T. 2014. Ralstonia
855 solanacearum type III secretion system effector Rip36 induces a
856 hypersensitive response in the nonhost wild eggplant Solanum torvum.
857 Molecular Plant Pathology 15(3):297-303 doi:10.1111/mpp.12079.
858 49. Parkinson, N., Bryant, R., Bew, J., Conyers, C., Stones, R., Alcock, M., and
859 Elphinstone, J. 2013. Application of Variable-Number Tandem-Repeat Typing
860 To Discriminate Ralstonia solanacearum Strains Associated with English
861 Watercourses and Disease Outbreaks. Applied and Environmental
862 Microbiology 79(19):6016-6022 doi:10.1128/aem.01219-13.

50. Peeters, N., Carrere, S., Anisimova, M., Plener, L., Cazale, A.-C., and Genin, S. 2013. Repertoire, unified nomenclature and evolution of the Type III effector gene set in the Ralstonia solanacearum species complex. BMC Genomics 14(1):859.

51. Poueymiro, M., Cunnac, S., Barberis, P., Deslandes, L., Peeters, N., Cazale-Noel, A. C., Boucher, C., and Genin, S. 2009. Two type III secretion system effectors from Ralstonia solanacearum GMI1000 determine host-range specificity on tobacco. Mol Plant Microbe Interact 22(5):538-50 doi:10.1094/MPMI-22-5-0538.

52. Poussier, S., Trigalet-Demery, D., Vandewalle, P., Goffinet, B., Luisetti, J., and Trigalet, A. 2000. Genetic diversity of Ralstonia solanacearum as assessed by PCR-RFLP of the hrp gene region, AFLP and 16S rRNA sequence analysis, and identification of an African subdivision. Microbiology 146(7):1679-1692.

53. Siri, M. I., Galvan, G. A., Quirici, L., Silvera, E., Villanueva, P., Ferreira, F., Franco Fraguas, L., and Pianzzola, M. J. 2009. Molecular marker diversity and bacterial wilt resistance in wild *Solanum commersonii* accessions from Uruguay. Euphytica 165:371-382.

54. Smith, J. J., Offord, L. C., Holderness, M., and Saddler, G. S. 1995. Genetic diversity of *Burkholderia solanacearum* (synonym *Pseudomonas solanacearum*) race 3 in Kenya. Applied and Environmental Microbiology 61(12):4263-4268.

55. Sokurenko, E. V., Gomulkiewicz, R., and Dykhuizen, D. E. 2006. Source-sink dynamics of virulence evolution. Nat Rev Micro 4(7):548-555.

887 56. Stajich, J. E., Block, D., Boulez, K., Brenner, S. E., Chervitz, S. A.,
888 Dagdigian, C., Fuellen, G., Gilbert, J. G. R., Korf, I., Lapp, H., Lehväslaiho,
889 H., Matsalla, C., Mungall, C. J., Osborne, B. I., Pocock, M. R., Schattner, P.,
890 Senger, M., Stein, L. D., Stupka, E., Wilkinson, M. D., and Birney, E. 2002.
891 The Bioperl Toolkit: Perl Modules for the Life Sciences. Genome Research
892 12(10):1611-1618 doi:10.1101/gr.361602.
893 57. Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and
894 post-analysis of large phylogenies. Bioinformatics 30(9):1312-1313
895 doi:10.1093/bioinformatics/btu033.
896 58. Stevens, P., and van Elsas, J. 2010. Genetic and phenotypic diversity of
897 Ralstonia solanacearum biovar 2 strains obtained from Dutch waterways.
898 Antonie van Leeuwenhoek 97(2):171-188 doi:10.1007/s10482-009-9400-1.
899 59. Strange, R. N., and Scott, P. R. 2005. Plant Disease: A Threat to Global Food
900 Security. Annual Review of Phytopathology 43(1):83-116
901 doi:doi:10.1146/annurev.phyto.43.113004.133839.
902 60. Stukenbrock, E. H., and McDonald, B. A. 2008. The Origins of Plant
903 Pathogens in Agro-Ecosystems. Annual Review of Phytopathology 46(1):75-
904 100 doi:doi:10.1146/annurev.phyto.010708.154114.
905 61. Tai, T. H., Dahlbeck, D., Clark, E. T., Gajiwala, P., Pasion, R., Whalen, M. C.,
906 Stall, R. E., and Staskawicz, B. J. 1999. Expression of the Bs2 pepper gene
907 confers resistance to bacterial spot disease in tomato. Proceedings of the
908 National Academy of Sciences 96(24):14153-14158
909 doi:10.1073/pnas.96.24.14153.
910 62. Timms-Wilson, T. M., Bryant, K., and Bailey, M. J. 2001. Strain
911 characterization and 16S-23S probe development for differentiating

geographically dispersed isolates of the phytopathogen Ralstonia solanacearum. Environmental Microbiology 3(12):785-797 doi:10.1046/j.1462-2920.2001.00250.x.

63. Trigalet, A., Caffier, D., and Gokelaere, T. 2002. Potato bacterial wilt [*Ralstonia solanacearum*]: 1. maintaining disease surveillance. - Pourriture brune "de la pomme de terre": 1. maintenir la surveillance. Pomme de Terre Francaise:45-8.

64. van Elsas, J. D., Kastelein, P., van Bekkum, P., van der Wolf, J. M., de Vries, P. M., and van Overbeek, L. S. 2000. Survival of Ralstonia solanacearum Biovar 2, the Causative Agent of Potato Brown Rot, in Field and Microcosm Soils in Temperate Climates. Phytopathology 90(12):1358-1366 doi:10.1094/PHYTO.2000.90.12.1358.

65. Vinatzer, B. A., Monteil, C. L., and Clarke, C. R. 2014. Harnessing Population Genomics to Understand How Bacterial Pathogens Emerge, Adapt to Crop Hosts, and Disseminate. Annual Review of Phytopathology 52(1):19-43 doi:doi:10.1146/annurev-phyto-102313-045907.

66. Vinatzer, B. A., Teitzel, G. M., Lee, M. W., Jelenska, J., Hotton, S., Fairfax, K., Jenrette, J., and Greenberg, J. T. 2006. The type III effector repertoire of Pseudomonas syringae pv. syringae B728a and its role in survival and disease on host and non-host plants. Mol Microbiol 62(1):26-44.

67. Wang, J.-F., Olivier, J., Thoquet, P., Mangin, B., Sauviac, L., and Grimsley, N. H. 2000. Resistance of Tomato Line Hawaii7996 to Ralstonia solanacearum Pss4 in Taiwan Is Controlled Mainly by a Major Strain-Specific Locus. Molecular Plant-Microbe Interactions 13(1):6-13 doi:10.1094/MPMI.2000.13.1.6.

68. Wang, J.-F., Ho, F.-I., Truong, H., Huang, S.-M., Balatero, C., Dittapongpitch, V., and Hidayati, N. 2013. Identification of major QTLs associated with stable resistance of tomato cultivar 'Hawaii 7996' to Ralstonia solanacearum. Euphytica 190(2):241-252 doi:10.1007/s10681-012-0830-x.

69. Wicker, E., Lefeuvre, P., de Cambiaire, J.-C., Lemaire, C., Poussier, S., and Prior, P. 2012. Contrasting recombination patterns and demographic histories of the plant pathogen Ralstonia solanacearum inferred from MLSA. ISME J 6(5):961-974 doi:http://www.nature.com/ismej/journal/v6/n5/suppinfo/ismej2011160s1.html .

70. Williamson, L., Nakaho, K., Hudelson, B., and Allen, C. 2002. *Ralstonia solanacearum* Race 3, Biovar 2 Strains Isolated from Geranium Are Pathogenic on Potato. Plant Disease 86(9):987-991 doi:10.1094/PDIS.2002.86.9.987.

71. Wroblewski, T., Caldwell, K. S., Piskurewicz, U., Cavanaugh, K. A., Xu, H., Kozik, A., Ochoa, O., McHale, L. K., Lahre, K., Jelenska, J., Castillo, J. A., Blumenthal, D., Vinatzer, B. A., Greenberg, J. T., and Michelmore, R. W. 2009. Comparative large-scale analysis of interactions between several crop species and the effector repertoires from multiple pathovars of pseudomonas and ralstonia. Plant Physiol 150(4):1733-49 doi:pp.109.140251 [pii]10.1104/pp.109.140251.

72. Xi'ou, X., Bihao, C., Guannan, L., Jianjun, L., Qinghua, C., Jin, J., and Yujing, C. 2014. Functional Characterization of a Putative Bacterial Wilt Resistance Gene (RE-bw) in Eggplant. Plant Molecular Biology Reporter:1-16 doi:10.1007/s11105-014-0814-1.

73. Xu, J., Zheng, H.-j., Liu, L., Pan, Z.-c., Prior, P., Tang, B., Xu, J.-s., Zhang, H., Tian, Q., Zhang, L.-q., and Feng, J. 2011. Complete Genome Sequence of the Plant Pathogen Ralstonia solanacearum Strain Po82. Journal of Bacteriology 193(16):4261-4262 doi:10.1128/jb.05384-11.